



National Data Warehouse

An Overview

- Purpose
- System Design
- NDW1 Project Scope
- User Benefits
- ➔ ■ **Data Integrity**
- Timeline
- Accessing Data in the NDW
- The Future



NDW Data Integrity

We are committed to assuring the integrity of your data throughout the NDW system.



What is “Data Integrity”

"Data Integrity" means that data have not been altered or destroyed in an unauthorized manner... [It] is both a security and a quality principle that prevents information from being modified or otherwise corrupted, either maliciously or accidentally."

National Committee on Vital and Health Statistics. Report to the Secretary of the U.S. Department of Health and Human Services on Uniform Data Standards for Patients Medical Record Information. July 2000



What is “Data Integrity”

... so, in simple terms, Data Integrity means that we will not make your good data bad, lose records, lose a diagnosis code, and so forth.

We should be able to show you that what we have at the NDW is exactly what you sent us.*

* Except for minor exceptions, e.g., non-Y2K compliant dates



What It is not...

Data integrity does not mean we can take poor quality data and change it to make it "good quality."

For example, data that is erroneous, missing, improperly coded, incomplete, etc. in the source systems will still be erroneous, missing, improperly coded, incomplete, etc. once it is loaded in the NDW.



"Data Quality"

The accuracy, completeness, appropriate use of codes, etc. of data in the source systems, its "data quality," has to be the responsibility of the local facilities who own and enter the data into their IT systems.

However we are partners. The NDW will provide local facilities information to assist them to improve the quality of the data in their systems.



NDW Data Quality Reports

- Field content checks against standard code tables
- Historical “norms” and deviation from norms
 - Records received versus expected
- Important missing data (e.g., SSNs)
- Timeliness of data



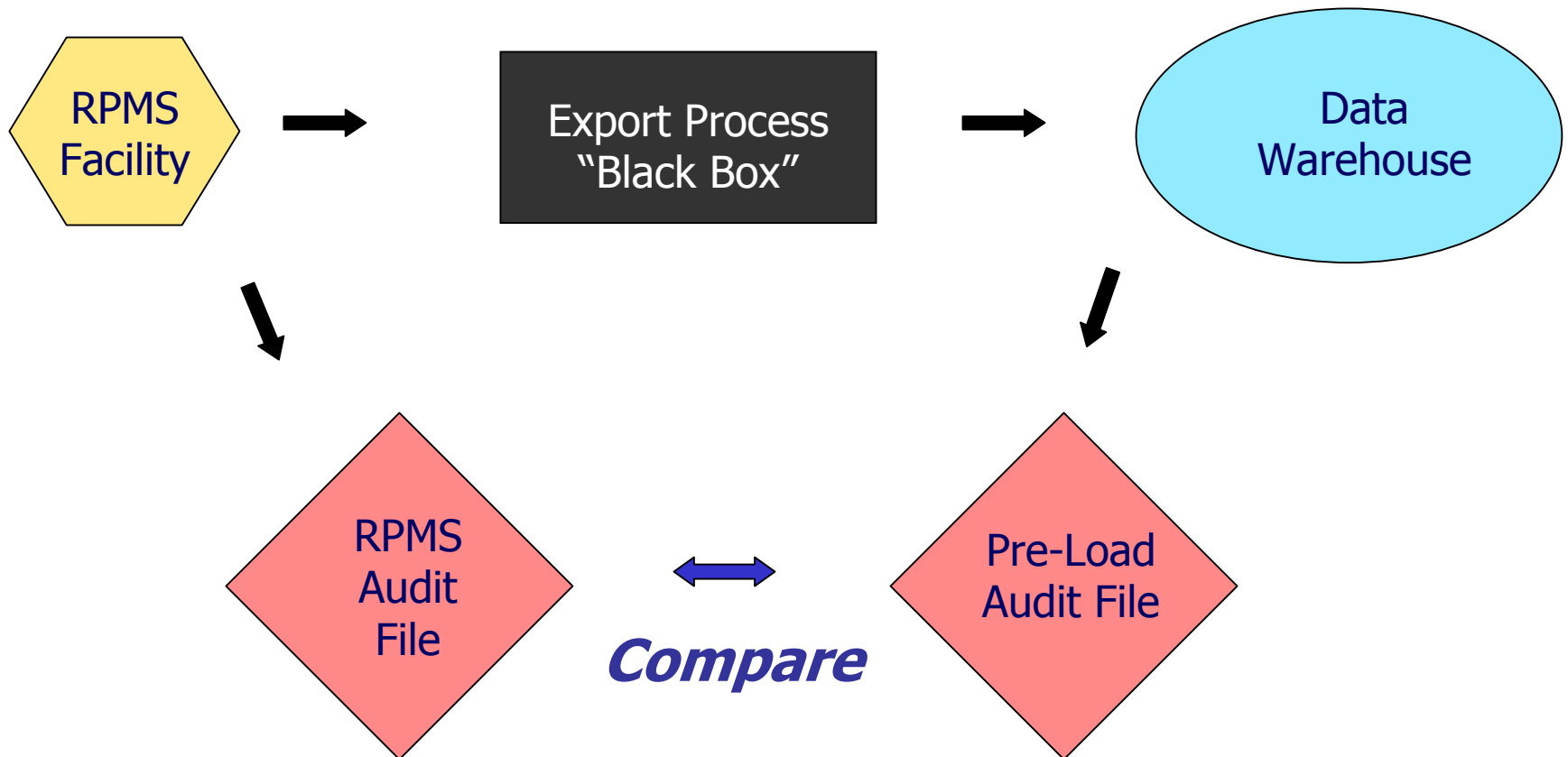
How Does the NDW Verify Data Integrity?

The NDW rigorously evaluates and monitors data integrity throughout the export of data, its routing to the NDW, and the load processes.

It does this in two steps:

- *one that monitors the entire export and routing process; and*
- *a second that monitors the load processes.*

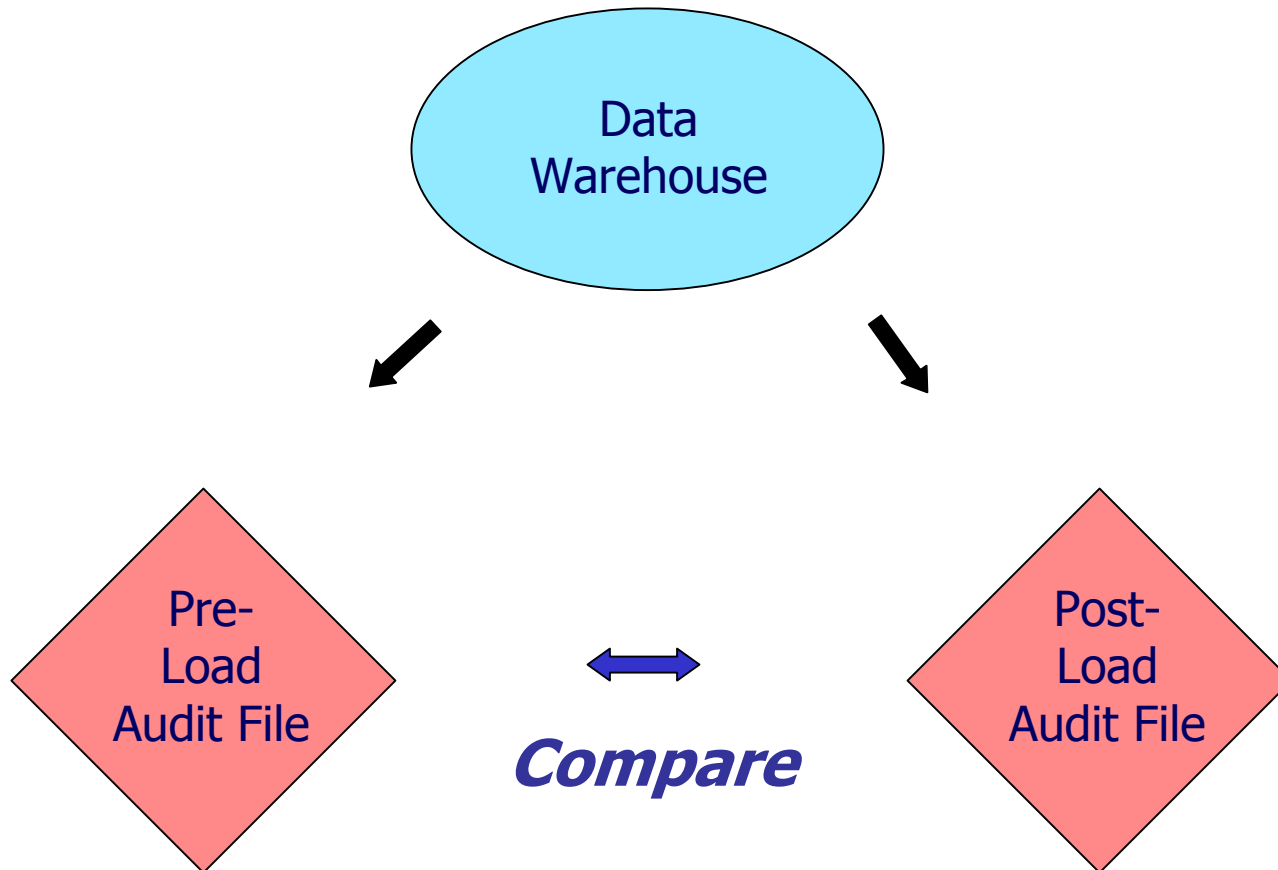
Data Integrity Verification Process – Step #1





Data Integrity Verification Process – Step #1

- At present the only export process we can monitor is IHS's RPMS export process
- But if non-RPMS sites send us appropriate audit files from their systems, we can do the same for them
- So we can even assist those with non-RPMS systems as they work to assure the integrity of their data through their proprietary exports!





What does the NDW Data Integrity Verification Process Monitor?

- Numbers of records
- Truncation of field entries
- Appropriate “mapping” of data elements from the source systems to the right place in the NDW system
- Serious errors that represent data corruption or processing errors



Summary Reports

- As each file export is received, we send an acknowledgement to the site
- After each file is loaded, we send a post-load report to the site
- After an Area's load is complete, we send them a summary report
- These will allow our customers to notify us about any discrepancies they see



Sample Acknowledgement

Sample of NDW Data Transmission Log Report

* DATA WAREHOUSE *
NDW Data Transmission Log Report
Data Received Date: Aug 25, 2004

Every file that
arrives at the
warehouse is
assigned a unique
Export ID number.

For more information please send an email to: ITSCDataWarehouse@ihs.gov

Export ID: 123
Export file name: 9999993040821090847.BDW
Static ASUFAC of exporting box: 999999
Beginning Date: Oct 01, 2000
Ending Date: Mar 31, 2001
Run Location: FACILITY NAME
Transmission Status: No errors detected
Total Number of Encounter Records Received: 0
Total Number of Registration Records Received: 99,999



Summary Reports

- As each file export is received, we send an acknowledgement to the site
- After each file is loaded, we send a post-load report to the site
- After an Area's load is complete, we send them a summary report
- These will allow our customers to notify us about any discrepancies they see

Sample Post-load Report

Sample of NDW Data Load Exception Report

* DATA WAREHOUSE *

NDW Post Data Load Report

Load Date: Aug 25, 2004

For more information please send an email to: ITSCDataWarehouse@hhs.gov

For error details, please visit: [www.](#)

A web site is being developed to provide additional information about the load exceptions.

Some Export ID number as reported on the acknowledgement.

Export ID: 123

Export file name: 9999993040821090847.L

Static ASUFAC of exporting box: 999999

Beginning Date: Oct 01, 2000

Ending Date: Mar 31, 2001

Run Location: FACILITY NAME

Load Status: Exceptions Detected

Total Number of Encounter Records Received: 0

Total Number of Registration Records Received: 99,999

Encounters:

Add	Change	Delete	Rejected
99,999	99,999	99,999	99,999

Registrations:

Add	Change	Delete	Rejected
99,999	99,999	99,999	99,999

Error Description	Field Name	Count
*** TRANSFORMATIONS ***		
INVALID DATE	Date Moved To Community (character format)	4,567
INVALID DATE	Date of Death (character format)	53
INVALID DATE	Eligibility Start Date (character format)	1,234
STANDARDIZING GENDER	Gender	1



Summary Reports

- As each file export is received, we send an acknowledgement to the site
- After each file is loaded, we send a post-load report to the site
- **After an Area's load is complete, we send them a summary report**
- These will allow our customers to notify us about any discrepancies they see



Sample Area Summary Report

Strawman sample of Encounter Summary Report

Area: XXXXXXXXXXXXXXXX

Service Type / Category

Month/YR -----	Added -----	Changed -----	Deleted -----	Rejected -----
X - XXXXXXXXXXXXXXXX / X - XXXXXXXXXXXXXXXX				
TOTAL	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
.
.
X - XXXXXXXXXXXXXXXX / X - XXXXXXXXXXXXXXXX				
TOTAL	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
.
.
X - XXXXXXXXXXXXXXXX / X - XXXXXXXXXXXXXXXX				
TOTAL	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
MMM YYYY	9999	9999	9999	9999
.
.



Summary Reports

- As each file export is received, we send an acknowledgement to the site
- After each file is loaded, we send a post-load report to the site
- After an Area's load is complete, we send them a summary report
- These will allow our customers to notify us about any discrepancies they see